

R. Swartz

1641

PAGE: 1

**RAW SEQUENCE LISTING
PATENT APPLICATION US/09/196,447**

DATE: 10/18/2000
TIME: 04:49:03

INPUT SET: S36023.raw

This Raw Listing contains the General Information Section and up to the first 5 pages.

SEQUENCE LISTING

3 (1) General Information:

ENTERED

9 (ii) TITLE OF INVENTION: NOVEL PARASITIC HELMINTH
10 P22U PROTEINS

12 (iii) NUMBER OF SEQUENCES: 17

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Sheridan Ross P.C.
16 (B) STREET: 1700 Lincoln St., Suite 3500
17 (C) CITY: Denver
18 (D) STATE: CO
19 (E) COUNTRY: U.S.A.
20 (F) ZIP: 80203

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Floppy disk
24 (B) COMPUTER: IBM PC compatible
25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

28 (vi) CURRENT APPLICATION DATA:

29 (A) APPLICATION NUMBER: US 09/196,447
30 (B) FILING DATE: 19-NOV-1998
31 (C) CLASSIFICATION:

33 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: US/08/460,428A
35 (B) FILING DATE: 02-JUN-1995

37 (viii) ATTORNEY/AGENT INFORMATION:

38 (A) NAME: Connell, Gary J.
39 (B) REGISTRATION NUMBER: 32,020
40 (C) REFERENCE/DOCKET NUMBER: 2618-13-3

42 (ix) TELECOMMUNICATION INFORMATION:

43 (A) TELEPHONE: 303/863-9700
44 (B) TELEFAX: 303/863-0223

RAW SEQUENCE LISTING
PATENT APPLICATION US/09/196,447DATE: 10/18/2000
TIME: 04:49:03

INPUT SET: S36023.raw

47 (2) INFORMATION FOR SEQ ID NO:1:

48

49 (i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 913 base pairs
51 (B) TYPE: nucleic acid
52 (C) STRANDEDNESS: single
53 (D) TOPOLOGY: linear

54

55 (ii) MOLECULE TYPE: cDNA

56

57 (ix) FEATURE:

58 (A) NAME/KEY: CDS
59 (B) LOCATION: 3..911

60

61

62 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

63

64 GC GAG TTA AAT AGT CGA ATT TCC GGA GTA CAC CGT AAT ACT GCA GGT 47

65 Glu Leu Asn Ser Arg Ile Ser Gly Val His Arg Asn Thr Ala Gly

66 1 5 10 15

67

68 GCT TTA CAA CGA TTT GCT CTA AAT GGT CAA AAT ACT CTT AAC GAA GGA 95

69 Ala Leu Gln Arg Phe Ala Leu Asn Gly Gln Asn Thr Leu Asn Glu Gly

70 20 25 30

71

72 TCA AGT TAT GAG CCA AAC GGA CTA TTT GTA TTT TCA GCA ATA AAC GGT 143

73 Ser Ser Tyr Glu Pro Asn Gly Leu Phe Val Phe Ser Ala Ile Asn Gly

74 35 40 45

75

76 AGC CAT ACT GAT AGC TTA TCT CAG TAT GGT GAA GGA ATA AAT GAA AAT 191

77 Ser His Thr Asp Ser Leu Ser Gln Tyr Gly Glu Gly Ile Asn Glu Asn

78 50 55 60

79

80 TAT CAT TCT GGA ACT AAT TAT TAT GAT GAA GTA GAA TTA AGA GAT AAA 239

81 Tyr His Ser Gly Thr Asn Tyr Tyr Asp Glu Val Glu Leu Arg Asp Lys

82 65 70 75

83

84 ACA AAT CAG ACA TCG TAC ATT AAT GGA AAT GAT AAT GGA ATC AAT GGA 287

85 Thr Asn Gln Thr Ser Tyr Ile Asn Gly Asn Asp Asn Gly Ile Asn Gly

86 80 85 90 95

87

88 AAG GAT GAT GAA GAT CTG GAT GAA TGC TCT GAT CAA GAA TTC CGA TGT 335

89 Lys Asp Asp Glu Asp Leu Asp Glu Cys Ser Asp Gln Glu Phe Arg Cys

90 100 105 110

91

92 CCA TAT CTA GCT AAA ACA CTT TGT GTT CAT TAT TTG AAA ATA TGC GAT 383

93 Pro Tyr Leu Ala Lys Thr Leu Cys Val His Tyr Leu Lys Ile Cys Asp

94 115 120 125

95

96 GGT ATT GAT GAT TGT GGT GAT GGA AGT GAT GAA ATG AAC TGT GCT GAT 431

97 Gly Ile Asp Asp Cys Gly Asp Gly Ser Asp Glu Met Asn Cys Ala Asp

98 130 135 140

99

**RAW SEQUENCE LISTING
PATENT APPLICATION US/09/196,447**

DATE: 10/18/2000
TIME: 04:49:03

INPUT SET: S36023.raw

100	GAT GAA GTG ATA ACA TCA ATA AAT GGT AAC GAA TCA ATC AAT ATC AGA	479
101	Asp Glu Val Ile Thr Ser Ile Asn Gly Asn Glu Ser Ile Asn Ile Arg	
102	145 150 155	
103		
104	TGT GAT CCG GAT CAA TTT CGA TGT GAA AAT GGA AAA TGT ATC GCA CAA	527
105	Cys Asp Pro Asp Gln Phe Arg Cys Glu Asn Gly Lys Cys Ile Ala Gln	
106	160 165 170 175	
107		
108	ATT GAT CGA TGT AAT CGA AAA TAT GAT TGT GAT GAT GGT ACA GAT GAA	575
109	Ile Asp Arg Cys Asn Arg Lys Tyr Asp Cys Asp Asp Gly Thr Asp Glu	
110	180 185 190	
111		
112	ACA ACT TGT GAA TAT TTC GTG CAA GCT TTG CAA CAA GCG AGA GGT GTA	623
113	Thr Thr Cys Glu Tyr Phe Val Gln Ala Leu Gln Gln Ala Arg Gly Val	
114	195 200 205	
115		
116	ACG GTG CAG GAT AAT GCA ATT CGA GAT GAC GAG ATA CCA AAT TAT ACT	671
117	Thr Val Gln Asp Asn Ala Ile Arg Asp Asp Glu Ile Pro Asn Tyr Thr	
118	210 215 220	
119		
120	GTA TCC ATG GAA CAG AAA TAC GAT CAA GTA AAG GAA GAT AAG GAG CGG	719
121	Val Ser Met Glu Gln Lys Tyr Asp Gln Val Lys Glu Asp Lys Glu Arg	
122	225 230 235	
123		
124	CGA ATG CAA GAG GAG GAA CAG GAA AGG CTG AGA GAG TAC GAG GAA	767
125	Arg Met Gln Glu Glu Glu Gln Glu Arg Leu Arg Glu Tyr Glu Glu	
126	240 245 250 255	
127		
128	CAG ATA CAG GAA AAA TTG AGG CAG GAG GAA GAA AGA GAA CGG CAA GAA	815
129	Gln Ile Gln Glu Lys Leu Arg Gln Glu Glu Glu Arg Glu Arg Gln Glu	
130	260 265 270	
131		
132	CAG GAA AGA AGA CAA AAG GAA CGA GAA AGA ATG GAA CAA GAA AGG ATA	863
133	Gln Glu Arg Arg Gln Lys Glu Arg Glu Arg Met Glu Gln Glu Arg Ile	
134	275 280 285	
135		
136	AGA CAA GAA TAT GAT GAA AAG GAA AGA CAA AGG CAA TAT GCT GAA CAG	911
137	Arg Gln Glu Tyr Asp Glu Lys Glu Arg Gln Arg Gln Tyr Ala Glu Gln	
138	290 295 300	
139		
140	GC	913
141		
142		
143	(2) INFORMATION FOR SEQ ID NO:2:	
144		
145	(i) SEQUENCE CHARACTERISTICS:	
146	(A) LENGTH: 303 amino acids	
147	(B) TYPE: amino acid	
148	(D) TOPOLOGY: linear	
149		
150	(ii) MOLECULE TYPE: protein	
151		
152	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:	

**RAW SEQUENCE LISTING
PATENT APPLICATION US/09/196,447**DATE: 10/18/2000
TIME: 04:49:04**INPUT SET: S36023.raw**

153
154 Glu Leu Asn Ser Arg Ile Ser Gly Val His Arg Asn Thr Ala Gly Ala
155 1 5 10 15
156
157 Leu Gln Arg Phe Ala Leu Asn Gly Gln Asn Thr Leu Asn Glu Gly Ser
158 20 25 30
159
160 Ser Tyr Glu Pro Asn Gly Leu Phe Val Phe Ser Ala Ile Asn Gly Ser
161 35 40 45
162
163 His Thr Asp Ser Leu Ser Gln Tyr Gly Glu Gly Ile Asn Glu Asn Tyr
164 50 55 60
165
166 His Ser Gly Thr Asn Tyr Tyr Asp Glu Val Glu Leu Arg Asp Lys Thr
167 65 70 75 80
168
169 Asn Gln Thr Ser Tyr Ile Asn Gly Asn Asp Asn Gly Ile Asn Gly Lys
170 85 90 95
171
172 Asp Asp Glu Asp Leu Asp Glu Cys Ser Asp Gln Glu Phe Arg Cys Pro
173 100 105 110
174
175 Tyr Leu Ala Lys Thr Leu Cys Val His Tyr Leu Lys Ile Cys Asp Gly
176 115 120 125
177
178 Ile Asp Asp Cys Gly Asp Gly Ser Asp Glu Met Asn Cys Ala Asp Asp
179 130 135 140
180
181 Glu Val Ile Thr Ser Ile Asn Gly Asn Glu Ser Ile Asn Ile Arg Cys
182 145 150 155 160
183
184 Asp Pro Asp Gln Phe Arg Cys Glu Asn Gly Lys Cys Ile Ala Gln Ile
185 165 170 175
186
187 Asp Arg Cys Asn Arg Lys Tyr Asp Cys Asp Asp Gly Thr Asp Glu Thr
188 180 185 190
189
190 Thr Cys Glu Tyr Phe Val Gln Ala Leu Gln Gln Ala Arg Gly Val Thr
191 195 200 205
192
193 Val Gln Asp Asn Ala Ile Arg Asp Asp Glu Ile Pro Asn Tyr Thr Val
194 210 215 220
195
196 Ser Met Glu Gln Lys Tyr Asp Gln Val Lys Glu Asp Lys Glu Arg Arg
197 225 230 235 240
198
199 Met Gln Glu Glu Glu Gln Glu Arg Leu Arg Glu Tyr Glu Glu Gln
200 245 250 255
201
202 Ile Gln Glu Lys Leu Arg Gln Glu Glu Glu Arg Glu Arg Gln Glu Gln
203 260 265 270
204
205 Glu Arg Arg Gln Lys Glu Arg Glu Arg Met Glu Gln Glu Arg Ile Arg

**RAW SEQUENCE LISTING
PATENT APPLICATION US/09/196,447**

DATE: 10/18/2000
TIME: 04:49:04

INPUT SET: S36023.raw

206	275	280	285
207			
208	Gln Glu Tyr Asp Glu Lys Glu Arg Gln Arg Gln Tyr Ala Glu Gln		
209	290	295	300

212 (2) INFORMATION FOR SEQ ID NO:3:

214 (i) SEQUENCE CHARACTERISTICS:
215 (A) LENGTH: 1016 base pa
216 (B) TYPE: nucleic acid
217 (C) STRANDEDNESS: single
218 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

221
222 (ix) FEATURE:
223 (A) NAME/KEY: CDS
224 (B) LOCATION: 3..626

225
226 (ix) FEATURE:
227 (A) NAME/KEY: 3'UTR
228 (B) LOCATION: 627...1016

231 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION **US/09/196,447**

DATE: 10/18/2000
TIME: 04:49:05

INPUT SET: S36023.raw

Line

Error

Original Text